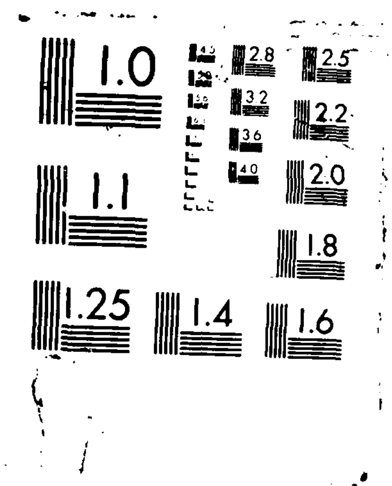
AD-A190 317 CONDITIONS FOR INFORMATION CAPACITY OF THE DISCRETE-TIME GAUSSIAN CHANNEL..(U) NORTH CAROLINA UNIV AT CHAPEL HILL DEPT OF STATISTICS C R BAKER DEC 87 1/1

UNCLASSIFIED LISS-22 N00014-86-K-0039 F/G 12/9 NL


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

Conditions for Information Capacity of the Discrete-Time Gaussian Channel to be Increased by Feedback*

LISS 22

December, 1987

Charles R. Baker

Department of Statistics

University of North Carolina

Chapel Hill, N.C. 27599

The results given here were partially presented at the IEEE International Symposium on Information Theory, Ann Arbor, Michigan, 5-8 October, 1986.

*Research Supported by ONR Contract #N00014-86-K-0039.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

ADA190317

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for Public Release: Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Departmet of Statistics 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code) University of North Carolina Chapel Hill, North Carolina 27514	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Office of Naval Research 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-86-K-0039
8c. ADDRESS (City, State and ZIP Code) Statistics & Probability Program Arlington, VA 22217	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT NO.
11. TITLE (Include Security Classification) Conditions for Information Capacity of the	
12. PERSONAL AUTHOR(S) C.R. Baker	
13a. TYPE OF REPORT TECHNICAL 13b. TIME COVERED FROM ______ TO ______	14. DATE OF REPORT (Yr., Mo., Day) Dec. 1987 15. PAGE COUNT 22
16. SUPPLEMENTARY NOTATION	
17. COSATI CODES FIELD / GROUP / SUB. GR.	18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) Channel capacity; Shannon theory; Information Theory; Channels with Feedback; Gaussian channels

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Sufficient conditions are given for optimal causal feedback to increase information capacity for the discrete-time additive Gaussian channel. The conditions are obtained by assuming linear feedback and reformulating the problem into an equivalent no-feedback problem.

TITLE CONT.: Discrete-Time Gaussian Channel to be Increased by Feedback

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED XX SAME AS RPT. XX DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION
22a. NAME OF RESPONSIBLE INDIVIDUAL C.R. Baker	22b. TELEPHONE NUMBER (Include Area Code) (919) 962-2189 22c. OFFICE SYMBOL

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

Abstract

Sufficient conditions are given for optimum causal feedback to increase information capacity for the discrete-time additive Gaussian channel. The conditions are obtained by assuming linear feedback and reformulating the problem into an equivalent no-feedback problem.

Introduction

Information capacity of the discrete-time additive Gaussian channel with feedback is an open problem. It has long been speculated that causal feedback can increase capacity. We give here sufficient conditions for optimum causal linear feedback to increase information capacity for any fixed value of the constraint, for all values of the constraint, and for all sufficiently large values of the constraint.

A special case of these results is for the finite-dimensional channel with a pure power constraint. The method developed here gives the solution for that case in a particularly easy fashion; see [1].

Recent work on the capacity of feedback channels has been done by Ihara [2] (for the finite-dimensional channel) and by Cover and Pombra [3].

Problem Statement

The capacity problem will be considered for both the infinite-dimensional and the finite-dimensional discrete-time additive Gaussian channel. However, the setup will be given only for ℓ_2; it will be seen (by substituting $\mathbb{R}^K$ for ℓ_2) that the procedure also applies without change to $\mathbb{R}^K$, although of course the finite-dimensional channel is much simpler and does not require the full development given here.

All stochastic processes are defined on a probability space (Ω,β,μ); $\underset{\sim}{E}(\cdot)$ will denote expectation with respect to μ. $\|x\|$ will denote the ℓ_2 norm of the vector x: $\|x\|^2 = \sum_{n\geq 1} [x(n)]^2$.

The channel output is $Y = X - BY + N$, where N is additive zero-mean Gaussian noise with strictly-positive trace-class covariance matrix R_N. X is a

message process independent of N, and B is a Hilbert-Schmidt strictly-lower-triangular (HSSLT) matrix: $\sum_{i,j\geq 1} b_{ij}^2 < \infty$ and $b_{ij} = 0$ for $j \geq i$, all $i \geq 1$. The mutual information of interest is that between X and Y, denoted I(X, Y). The constraint will be given in terms of a trace-class covariance matrix R_W. Any constraint must imply a constraint of this form if the capacity is to be finite [4]. The class of admissible message processes X and HSSLT matrices B consist of all X such that almost all sample paths of X belong to range($R_W^{\frac{1}{2}}$), range(B) is contained in range($R_W^{\frac{1}{2}}$), and $\underset{\sim}{E}\|X-BY\|_W^2 \leq P$, where $\|u\|_W^2 = \|R_W^{-\frac{1}{2}}u\|^2$. The capacity is then the supremum of the mutual information I(X,Y) over all such admissible pairs (X,B).

The feedback capacity will be denoted by $C_{WF}(P)$. The capacity of this channel without feedback is for the case B = 0, so that the constraint is $\underset{\sim}{E}\|X\|_W^2 \leq P$. This capacity will be denoted by $C_W(P)$.

The assumptions that R_N and R_W are strictly-positive can be dropped. Attention can be restricted to $\overline{\text{range}(R_N)}$; in order to have finite capacity, one must then have that R_W is strictly positive as an operator in $\overline{\text{range}(R_N)}$; see [4] for details. However, without loss of generality, it is assumed here that both R_W and R_N are strictly positive.

Preliminaries

This section contains several mathematical definitions and small results that will be needed to prove the main result. It will be seen that much of this is obvious when one treats $\mathbb{R}^K$.

$\ell_2 \otimes \ell_2$ will denote the set of all Hilbert-Schmidt operators mapping ℓ_2 into ℓ_2. A is in $\ell_2 \otimes \ell_2$ if and only if A has a matrix representation such that

$\sum_{i,j\geq 1} [A(ij)]^2 < \infty$. For A_1 and A_2 in $\ell_2 \otimes \ell_2$, define

$$\langle A_1, A_2 \rangle_{\otimes} = \text{Trace } A_1 A_2^{*}$$
$$= \sum_i \sum_j A_1(ij) A_2^{*}(ji).$$

$\langle \cdot, \cdot \rangle_{\otimes}$ defines an inner product on $\ell_2 \otimes \ell_2$, and it is known that $\ell_2 \otimes \ell_2$ is a Hilbert space under this inner product [5]. Moreover, convergence of a sequence (A_n) in $\ell_2 \otimes \ell_2$ to an element A in $\ell_2 \otimes \ell_2$ implies that $\|A_n - A\| \to 0$ and thus $A_n x \to Ax$ for all x in ℓ_2.

(δ_n) will denote the natural basis vectors in ℓ_2: $\delta_n(i) = 0$ for $i \neq n$, $\delta_n(n) = 1$. Let $H_n = \text{span}\{\delta_i, i \leq n\}$ and denote by P_n the projection operator with range space H_n. P_n is a diagonal matrix with $P_n(i,i) = 1$ for $i \leq n$; $P_n(i,i) = 0$ for $i > n$.

<u>Lemma 1</u>: Let R_W^n be the matrix $P_n R_W P_n$:

$$R_W^n(ij) = R_W(ij) \text{ for } i \leq n,\ j \leq n$$
$$= 0 \text{ otherwise. Then:}$$

(1) $R_W^n = V_n V_n^{*}$ for a lower triangular matrix V_n with $V_n(ii) = c_i$ for $i \leq n$, where $\prod_{i=1}^{n} c_i^2 = \text{determinant } R_W^n$;

(2) $V_n x = 0$ for all x in $H_n^{\perp}$;

(3) If $m > n$, $V_n = P_n V_m$;

(4) (V_n) is a Cauchy sequence in $\ell_2 \otimes \ell_2$;

(5) $R_W = VV^{*}$, where V is a unique lower-triangular Hilbert-Schmidt matrix such that $V(i,i) = c_i$ for all $i \geq 1$, and $V = \lim V_n$ in the topology of $\ell_2 \otimes \ell_2$. Moreover, $V_n = P_n V$ for $n \geq 1$.

<u>Proof</u>: Since H_n can be identified with $\mathbb{R}^n$, and R_W^n with a covariance matrix in $\mathbb{R}^n$, (1) is obvious, as is (2). To see (3), if $i,j \leq n$ and $m > n$, then

$$R_W^n(ij) = R_W^m(ij) = \sum_{k\le n} V_m(ik)V_m(jk)$$

$$= \sum_{k\le \min(i,j)} [P_nV_m](ik)[P_nV_m](jk).$$

Since P_nV_m is lower triangular, $P_nV_m(ii) = V_m(ii) = V_n(ii)$ for all $i \le n$, and the factorization $R_W^n = V_nV_n^*$ is unique when the diagonal elements of V_n are fixed [7], it follows that $V_n = P_nV_m$.

For (4), note that $R_W^n = P_nR_WP_n$. If $m > n$, then

$\text{Trace } (V_n-V_m)(V_n-V_m)^* = \sum_{j\ge 1} \|(V_n-V_m)^*\delta_j\|^2 = \sum_{j=1}^{m} \|V_m^*(P_n-I)\delta_j\|^2 = \sum_{j=n+1}^{m} \|V_m^*\delta_j\|^2 =$ $\sum_{j=n+1}^{m} \langle P_mR_WP_m\delta_j,\delta_j\rangle = \sum_{j=n+1}^{m} \langle R_W\delta_j,\delta_j\rangle$. Since R_W has finite trace, this sum converges to zero as $m,n \to \infty$, showing that (V_n) is Cauchy in $\ell_2\otimes\ell_2$.

To obtain (5), we first recall that convergence in $\ell_2\otimes\ell_2$ implies norm convergence to the same limit [5], so there exists by (4) a Hilbert-Schmidt operator V such that $V_n \to V$ in both $\ell_2\otimes\ell_2$ and operator norm. $V_nV_n^*$ must then converge to VV^* in the operator norm topology. However, $V_nV_n^* = R_W^n = P_nR_WP_n$, so $V_nV_n^*$ converges to R_W in operator norm. Since the set of bounded linear operators on ℓ_2 is a Banach space under the operator norm [6], a Cauchy sequence has a unique limit, and this gives $R_W = VV^*$. V is necessarily Hilbert-Schmidt, since R_W is trace-class. To see that V must be lower-triangular, note that $\text{Tr } (V_n-V)(V_n-V)^* \to 0$, and $\text{Tr } (V_n-V)(V_n-V)^* =$ $\sum_{i,j} (V_n(ij)-V(ij))^2$. Since $V_n(ij) = 0$ for $j > i$ and all $n \ge 1$, it follows that $V(ij) = 0$ for $j > i \ge 1$. The same relations show that $V(ii) = c_i$ for all $i \ge 1$. Since $P_nVV^*P_n = V_nV_n^*$, we proceed as before to obtain $V_n = P_nV$.

□

If u and v are two vectors in ℓ_2, then $u\otimes v$ is defined to be the element of $\ell_2\otimes\ell_2$ defined by $(u\otimes v)x = \langle v,x\rangle u$.

In order that the capacity without feedback be finite, it is necessary and sufficient that $R_N = R_W^{\frac{1}{2}}(I+S)R_W^{\frac{1}{2}}$ where S is a self-adjoint operator in ℓ_2 such that $(I+S)^{-1}$ exists and is bounded [4]. The limit points of the spectrum of S consist of all real numbers λ such that λ is either an eigenvalue of S of infinite multiplicity, or the limit of a sequence of distinct eigenvalues, or a point of the continuous spectrum (i.e., $(S-\lambda I)^{-1}$ exists and is densely-defined but not bounded). The set of limit points of S is not empty. For discussion of these and related facts, see [6]. θ will be used to denote the smallest limit point of the spectrum of S. As in [4], (λ_n) will always be used to denote the sequence of eigenvalues of S that are strictly less than θ; they are ordered by $\lambda_1 \leq \lambda_2 \leq \ldots < \theta$, and repeated in the sequence according to their multiplicity. Of course, there may not be any eigenvalues strictly less than θ. If $\{\lambda_n, n \geq 1\}$ is not empty, then $\{e_n, n \geq 1\}$ will denote orthonormal eigenvectors of S corresponding to the eigenvalues (λ_n): $Se_n = \lambda_n e_n$, $n \geq 1$.

With $R_W = VV^*$, $R_W^{\frac{1}{2}} = VL^*$ for L a unitary operator [7]. Since $I + S = R_W^{-\frac{1}{2}}R_N R_W^{-\frac{1}{2}}$ (on the range of $R_W^{\frac{1}{2}}$), $L^*(I+S)L = I + L^*SL = V^{-1}R_N V^{*-1}$. As L is unitary, L^*SL has the same spectrum as S, and so $V^{-1}R_N V^{*-1}$ has the same spectrum as I + S. Thus, $1 + \theta$ is the smallest limit point of the spectrum of $V^{-1}R_N V^{*-1}$ and $\{1+\lambda_k, k \geq 1\}$ are the eigenvalues of $V^{-1}R_N V^{*-1}$ that are strictly less than $1 + \theta$.

Main Result

Theorem: Let V be a lower-triangular matrix such that $R_W = VV^*$. Fix $P > 0$.

(1) For the K-dimensional channel, let $\beta_1 \leq \beta_2 \leq \ldots \leq \beta_K$ be the eigenvalues of S, with J the largest integer $\leq K$ such that $J\beta_J \leq P + \sum_{i=1}^{J} \beta_i$.

$C_{WF}(P) > C_W(P)$ if the set of eigenvectors of $V^{-1}R_N V^{*-1}$ corresponding to the sequence of eigenvalues $(1 + \beta_i,\ i \le J)$ does not contain J natural basis vectors.

(2) For the infinite-dimensional channel, $C_{WF}(P) > C_W(P)$ if the following conditions are satisfied:

(a) $\{\lambda_k,\ k \ge 1\}$ is not empty;

(b) If there exists a largest integer J such that $J\lambda_J \le P + \sum_{i=1}^{J} \lambda_i$, then $C_{WF}(P) > C_W(P)$ if the set of eigenvectors of $V^{-1}R_N V^{*-1}$ corresponding to the sequence of eigenvalues $(1 + \lambda_k,\ k \le J)$ does not contain J natural basis vectors.

or (b') If $J\lambda_J \le P + \sum_{i=1}^{J} \lambda_i$ for all λ_J, then $C_{WF}(P) > C_W(P)$ if the subspace spanned by the eigenvectors of $V^{-1}R_N V^{*-1}$ corresponding to the sequence of eigenvalues $(1 + \lambda_k,\ k \ge 1)$ does not contain a set of natural basis vectors that is complete for that subspace and which are eigenvectors of $V^{-1}R_N V^{*-1}$.

<u>Corollary</u>:

(1) For the finite-dimensional discrete-time channel, causal linear feedback can increase information capacity for all $P > 0$ if the subspace spanned by the eigenvectors of $V^{-1}R_N V^{*-1}$ corresponding to its smallest eigenvalue does not contain a basis consisting entirely of natural basis vectors which are eigenvectors of $V^{-1}R_N V^{*-1}$. Causal linear feedback can increase capacity for all sufficiently large P if and only if $V^{-1}R_N V^{*-1}$ is not diagonal.

(2) For the infinite-dimensional discrete-time channel, causal linear feedback can increase capacity for all $P > 0$ if $\{\lambda_k,\ k \ge 1\}$ is not empty

and the subspace spanned by the eigenvectors of $V^{-1}R_NV^{*-1}$ corresponding to its smallest eigenvalue does not contain a basis consisting entirely of natural basis vectors which are eigenvectors of $V^{-1}R_NV^{*-1}$.

$C_{WF}(P) > C_W(P)$ for all sufficiently large P if $\{\lambda_k, k \geq 1\}$ is not empty and the subspace spanned by the eigenvectors corresponding to the eigenvalues $\{1 + \lambda_k, k \geq 1\}$ of $V^{-1}R_NV^{*-1}$ does not contain a basis for the subspace consisting entirely of natural basis vectors which are eigenvectors of $V^{-1}R_NV^{*-1}$.

Remark: The sufficient condition in (2) giving $C_{WF}(P) > C_W(P)$ for all sufficiently large P is equivalent to the following statement: $\{\lambda_k, k \geq 1\}$ is not empty, and the restriction of $V^{-1}R_NV^{*-1}$ to the subspace spanned by the eigenvectors of $V^{-1}R_NV^{*-1}$ corresponding to the eigenvalues $\{1 + \lambda_k, k \geq 1\}$ is not a diagonal matrix.

The results stated in (1) of the Corollary were proved in [1], where the development is much streamlined because of the simpler nature of the finite-dimensional problem. That work used $R_W = I$. For the same finite-dimensional channel and constraint, Ihara has obtained the result that capacity is increased for all sufficiently large P if R_N is not diagonal [2], although his result is stated in a different form; his methods are quite different from those used here. He also gives as a sufficient condition for $C_{WF}(P) > C_W(P)$ for all $P > 0$ the condition that (in the terminology used here) R_N has no natural basis vectors as eigenvectors. The corresponding sufficient condition given in (1) of the Corollary is much weaker.

Reformulation of the Problem

In this section, the original linear feedback problem is converted into an equivalent no-feedback problem. Originally, $Y = X - BY + N$, where the matrix B is HSSLT. $(I+B)^{-1}$ exists, since B can have no non-zero eigenvalues; $(I+B)^{-1}$ is bounded, since B is compact (and thus has only zero as a limit point of the spectrum). Thus, $Y = (I+B)^{-1}X + (I+B)^{-1}N$. Since $(I+B)^{-1}$ is 1:1,

$$I(X, X+N) = I(X, (I+B)^{-1}(X+N)) = I(X, Y).$$

Of course, the constraint $\underset{\sim}{E}\|X-BY\|_W^2 \leq P$ is the same as $\underset{\sim}{E}\|X-B(I+B)^{-1}(X+N)\|_W^2 \leq P$. Using $R_W = VV^*$ with V Hilbert-Schmidt and lower triangular, the constraint can be written

$$P \geq \underset{\sim}{E}\|V^{-1}X - V^{-1}B(I+B)^{-1}(X+N)\|^2$$

or,

$$P \geq \underset{\sim}{E}\|Z - D(V^{-1}Z+N)\|^2,$$

where $D = V^{-1}B(I+B)^{-1}$ and $Z = V^{-1}X$. D is well-defined and bounded, since range(B) $\subset$ range(V) and $(I+B)^{-1}$is bounded. Moreover, since B is HSSLT and both V^{-1} and $(I+B)^{-1}$ are lower triangular, D must be strictly lower triangular and bounded (BSLT).

The feedback capacity problem under our initial assumptions thus becomes

$$\text{maximize } I(X, X+N)$$

$$\text{subject to } P \geq \underset{\sim}{E}\|V^{-1}X - D(X+N)\|^2$$

where D is permitted to be any bounded SLT matrix in ℓ_2.

This is actually the problem that will be considered below in obtaining the sufficient conditions of the Theorem and the Corollary.

Analysis

Let $H(\ell_2,\mu)$ be the set of all real random vectors u on (Ω,β) such that $u(\omega) \in \ell_2$ a.e. $dP(\omega)$ and $\underset{\sim}{E} \sum_{n\geq 1} [u(n,\omega)]^2 < \infty$. $H(\ell_2,\mu)$ is a Hilbert space under the inner product $(f,g)_\mu = \underset{\sim}{E} \sum_{n\geq 1} f(n,\omega)g(n,\omega)$. Let Y_1 and Y_2 be two mutually independent zero-mean Gaussian random vectors in ℓ_2: $\underset{\sim}{E}\, Y_1(m,\omega)Y_2(n,\omega) = 0$ for all $n,m \geq 1$. Suppose that $R_{Y_1} + R_{Y_2}$ is strictly positive. Define $H_-(Y_1+Y_2)$ as the set of all elements f in $H(\ell_2,\mu)$ such that $f = B(Y_1+Y_2)$ for some bounded SLT matrix operator B. $H_-(Y_1+Y_2)$ is clearly a linear manifold in $H(\ell_2,\mu)$. To see that this linear manifold is closed, one notes that if (B^n) is a sequence of bounded SLT operators,

$$\|B^n(Y_1+Y_2) - B^m(Y_1+Y_2)\|_\mu^2 = \|(B^n-B^m)(Y_1+Y_2)\|_\mu^2$$

$$= \text{Trace } (B^n-B^m)(R_{Y_1}+R_{Y_2})(B^n-B^m)^*$$

$$\geq \|(B^n-B^m)(R_{Y_1}+R_{Y_2})^{\frac{1}{2}}\|^2 \geq \|B^n-B^m\|^2\gamma_0$$

where γ_0 is the smallest eigenvalue of $R_{Y_1} + R_{Y_2}$. Thus, $(B^n(Y_1+Y_2))$ Cauchy in $H(\ell_2,\mu)$ implies that (B^n) is Cauchy in operator norm, so converges to a bounded linear operator B. To see that B is SLT, one notes that $R_{Y_1} + R_{Y_2} = QQ^*$ for some lower-triangular Q, and so $(R_{Y_1}+R_{Y_2})^{\frac{1}{2}} = QT$ for T unitary [8]. This gives

$$\|B^n(Y_1+Y_2) - B^m(Y_1+Y_2)\|_\mu^2 = \|(B^n-B^m)(R_{Y_1}+R_{Y_2})^{\frac{1}{2}}\|_\otimes^2$$

$$= \|(B^n-B^m)Q\|_\otimes^2.$$

Thus, (B^nQ) is Cauchy in $\ell_2\otimes\ell_2$, so that BQ must be strictly lower triangular;

since Q^{-1} exists and is lower triangular, this shows that B is a bounded SLT matrix operator, so that $H_-(Y_1+Y_2)$ is closed.

Now consider our feedback problem: We wish to maximize $I(X, X+N)$ subject to $P \geq \underset{\sim}{E}\|V^{-1}X - D(X+N)\|^2 = \|V^{-1}X - D(X+N)\|_\mu^2$, where D is permitted to be any bounded SLT matrix. Given any choice of D that satisfies this constraint, we know that $\|V^{-1}X - D(X+N)\|_\mu^2 \geq \|V^{-1}X - P_-(V^{-1}X)\|_\mu^2$, where $P_-(V^{-1}X)$ is the projection of $V^{-1}X$ onto $H_-(X+N)$. Thus, we can assume WLOG that D is the optimum bounded SLT matrix for minimizing the distance in $H(\ell_2,\mu)$ norm between $V^{-1}X$ and $H_-(X+N)$; $D(X+N)$ is the projection of $V^{-1}X$ onto $H_-(X+N)$.

Now, let X be the optimum no-feedback message for the case when capacity is attained (assuming here that $C_W(P)$ can be attained). As the message for the feedback problem, use αX. Then $I(\alpha X, \alpha X+N) > C_W(P)$ if $\alpha > 1$. Choose α to satisfy the constraint:

$$\begin{aligned} P &= \|V^{-1}\alpha X - D(\alpha X+N)\|_\mu^2 \\ &= \alpha^2 \mathrm{Tr}\ V^{-1}R_X V^{-1} - \Delta, \end{aligned}$$

where $\Delta = \Delta(\alpha)$ is the $H(\ell_2,\mu)$ norm of the projection of $\alpha V^{-1}X$ on $H_-(\alpha X+N)$. Since $\mathrm{Tr}\ V^{-1}R_X V^{-1} = \mathrm{Tr}\ R_W^{-\frac{1}{2}}R_X R_W^{-\frac{1}{2}} = \underset{\sim}{E}\|X\|_W^2 = P$, we have $P = \alpha^2 P - \Delta$, so that $\alpha > 1$ if $\alpha V^{-1}X$ is not orthogonal to $H_-(\alpha X+N)$.

<u>Prop.</u>: $\alpha V^{-1}X$ is orthogonal to $H_-(\alpha X+N)$ if and only if $V^{-1}R_X V^{*-1}$ is diagonal.

<u>Proof</u>: Since X is independent of N, $(\alpha V^{-1}X, D(\alpha X+N))_\mu = \alpha^2 \mathrm{Tr}\ DR_X V^{*-1} = \alpha^2 \mathrm{Tr}\ DV[V^{-1}R_X V^{*-1}]$. If $V^{-1}R_X V^{*-1}$ is diagonal, then (as DV is SLT) $\mathrm{Tr}\ DR_X V^{*-1} = 0$ for every bounded SLT matrix D.

If $\mathrm{Tr}\ DR_X V^{*-1} = 0$ for every bounded SLT matrix D, take i,j with $i > j$.

Let $D(k\ell) = 0$ unless $k = i$, $\ell = j$;

$D(ij) = 1$.

Then, $\mathrm{Tr}\ DV(V^{-1}R_X V^{*-1}) = (R_X V^{*-1})(ji) = 0$. This shows that $R_X V^{*-1}$ must be lower triangular, so that $V^{-1}R_X V^{*-1}$ must also be lower triangular. Since $V^{-1}R_X V^{*-1}$ is symmetric, $V^{-1}R_X V^{*-1}$ must be diagonal. □

In the above development, we have implicitly assumed that α always exists to solve the equation $\alpha^2 P = P + \Delta(\alpha)$. This is not obvious, as the subspace $H_-(\alpha X+N)$ changes with α. Here we will show that a lower bound α_1 exists; i.e., $\alpha_1 > 1$ and a bounded SLT D such that $\|\alpha_1 V^{-1}X - D(\alpha_1 X+N)\|_\mu^2 = P$.

Let D be the optimum SLT matrix to minimize $\|V^{-1}X - D(X+N)\|_\mu^2$. Then $\mathrm{Tr}\ DR_X V^{*-1} = \mathrm{Tr}\ DR_X D^* + \mathrm{Tr}\ DR_N D^* = \Delta(1)$. Take $\alpha \neq 0$. Then $\|\alpha V^{-1}X - D(\alpha X+N)\|_\mu^* = \alpha^2(P - 2\mathrm{Tr}\ DR_X V^{*-1} + \mathrm{Tr}\ DR_X D^*) + \mathrm{Tr}\ DR_N D^*$. If $P - 2\mathrm{Tr}\ DR_X V^{*-1} + \mathrm{Tr}\ DR_X D^* \neq 0$, then one can set $P = \|\alpha V^{-1}X - D(\alpha X+N)\|_\mu^2$ and solve for α^2, obtaining

$$\alpha^2 = \frac{P - \mathrm{Tr}\ DR_N D^*}{P - \mathrm{Tr}\ DR_N D^* - \Delta(1)}, \text{ giving } \alpha > 1.$$

To see that $P - 2\mathrm{Tr}\ DR_X V^{*-1} + \mathrm{Tr}\ DR_X D^* \neq 0$ when $V^{-1}R_X V^{*-1}$ is not diagonal, we note that if inequality does not hold, then $\|V^{-1}X - D(X+N)\|_\mu^2 = P - \Delta(1) = \mathrm{Tr}\ DR_N D^*$. Similarly, for any $\alpha \neq 0$, $\|\alpha V^{-1}X - D(\alpha X+N)\|_\mu^2 = \alpha^2(P - 2\mathrm{Tr}\ DR_X V^{*-1} + \mathrm{Tr}\ DR_X D^*) + \mathrm{Tr}\ DR_N D^* = \mathrm{Tr}\ DR_N D^*$. Thus, $P - \Delta(\alpha) \leq \mathrm{Tr}\ DR_N D^* = P - \Delta(1)$, or $\Delta(\alpha) \geq \Delta(1)$, all $\alpha \neq 0$. This cannot hold if $V^{-1}R_X V^{*-1}$ is not diagonal, since $\Delta(\alpha) \leq \alpha^2 P$, and $\Delta(1) \neq 0$ when $V^{-1}R_X V^{*-1}$ is not diagonal.

We have now shown that causal linear feedback can increase capacity provided that $V^{-1}R_X V^{*-1}$ is not diagonal, where R_X is the optimum message covariance matrix in the no-feedback problem (whenever capacity can be attained in the no-feedback case). These conditions need to be converted into conditions on the noise covariance matrix R_N and the constraint matrix R_W.

This will be done in the next two sections, treating the $\mathbb{R}^K$ and ℓ_2 channels.

Finite-Dimensional Channel ($\mathbb{R}^K$)

From Theorem 1 of [4], the optimum no-feedback message has covariance matrix given by

$$R_X = \frac{1}{J}\left[\sum_{i=1}^{J}\beta_i + P\right]\sum_{n=1}^{J} R_W^{\frac{1}{2}}u_n \otimes R_W^{\frac{1}{2}}u_n - \sum_{m=1}^{J}\beta_m R_W^{\frac{1}{2}}u_n \otimes R_W^{\frac{1}{2}}u_n,$$

where $\{u_n, n \le K\}$ are o.n. eigenvectors of S corresponding to the increasing sequence of eigenvalues (β_n), and J is the largest integer $\le K$ such that $P + \sum_{i=1}^{J}\beta_i \ge J\beta_J$. Let L be the unitary operator in ℓ_2 such that $R_W^{\frac{1}{2}} = VL^*$. Then

$$V^{-1}R_X V^{*-1} = \frac{1}{J}\sum_{n=1}^{J}\left[\sum_{i=1}^{J}\beta_i + P - \beta_n\right](L^*u_n)\otimes(L^*u_n).$$

Now, $V^{-1}R_N V^{*-1} = L^*(I+S)L$, and $Su_n = \beta_n u_n$, so

$$L^*(I+S)LL^*u_n = L^*(I+S)u_n = (1+\beta_n)L^*u_n;$$

i.e., $\{L^*u_n, n \le K\}$ are c.o.n. eigenvectors of $V^{-1}R_N V^{*-1}$ corresponding to the sequence of eigenvalues $(1+\beta_n)$, $n \le K$. $V^{-1}R_X V^{*-1}$ is then diagonal if and only if $\{L^*u_n, n \le J\}$ can be taken as natural basis vectors, proving (1) of the Theorem. For all sufficiently small $P > 0$,

$$V^{-1}R_X V^{*-1} = \frac{P}{M}\sum_{n=1}^{M} L^*u_n \otimes L^*u_n,$$

where M is the multiplicity of the eigenvalue β_1 of S, and of the eigenvalue $1 + \beta_1$ of $V^{-1}R_N V^{*-1}$. Thus, $V^{-1}R_X V^{*-1}$ cannot be diagonal if $\{L^*u_n, n \le M\}$ cannot be taken as natural basis vectors. For larger values of P, when R_X has the representation given above for $J > M$, then the eigenvectors of $V^{-1}R_X V^{*-1}$ must include $\{L^*u_n, n \le M\}$. Now, if $V^{-1}R_X V^{*-1}$ is diagonal, then it must have a

c.o.n. set of eigenvectors consisting of natural basis vectors. However, $\mathrm{span}\{L^* u_n,\ n \le M\}$ cannot be spanned by M natural basis vectors, so that $V^{-1}R_X V^{*-1}$ cannot have a c.o.n. set of eigenvectors consisting entirely of natural basis vectors. This shows that $C_{WF}(P) > C_W(P)$ for every $P > 0$ if the M-dimensional eigenmanifold of β_1 is not spanned by M natural basis vectors.

By letting P become sufficiently large, J = K, and then the above expressions show that $V^{-1}R_X V^{*-1}$ will be diagonal if and only if $V^{-1}R_N V^{*-1}$ is diagonal: $V^{-1}R_X V^{*-1} = \frac{1}{K}\left[\sum_{i=1}^{K} \beta_i + P + K\right]I - V^{-1}R_N V^{*-1}$. This proves the sufficient conditions of the Corollary.

To see that capacity cannot be increased by causal feedback if $V^{-1}R_N V^{*-1}$ is diagonal, one notes that the feedback capacity problem is that of maximizing I(X, X+N) subject to the constraint $\underset{\sim}{E}\|V^{-1}D(X,N)\|_n^2 \le P$, where $\|x\|_n^2 \equiv \sum_{i=1}^{n} x_i^2$ and D is a possibly non-linear operator depending only on the past of the second coordinate (causal): $[D(x,y)]_n = D_1^n(x,[y_1,y_2,\ldots,y_{n-1}])$, where D_1^n maps $\mathbb{R}^K \times \mathbb{R}^{n-1}$ into $\mathbb{R}$. Write the constraint as $\underset{\sim}{E}\|V^{-1}F(T,Z)\|_n^2 \le P$, where $T = V^{-1}X$, $Z = V^{-1}N$, and $F(x,y) = D[Vx,Vy]$. Since V is lower-triangular and D is causal in the second coordinate, F is also causal in the second coordinate. $I(X,Y) = I(V^{-1}X,\ V^{-1}Y) = I[V^{-1}X,\ V^{-1}D(X,N) + V^{-1}N] = I[T,\ V^{-1}F(T,Z) + Z]$. The constraint is $E\|V^{-1}F(T,Z)\|_n^2 \le P$, and $V^{-1}F(x,y)$ is a causal function of y. However, Z has covariance matrix $V^{-1}R_N V^{*-1}$. Thus, if $V^{-1}R_N V^{*-1}$ is diagonal, the original problem is equivalent to the capacity problem with causal feedback when the channel is without memory. It is well-known that capacity cannot be increased in this case.

This completes the proof of the theorem and corollary for the finite-dimensional channel.

Infinite-Dimensional Channel (ℓ_2)

First, assume that $\{\lambda_n, n \geq 1\}$ is not empty. Several cases need to be considered. The various expressions for the optimum R_X (when it exists) and the value of $C_W(P)$ are taken from [4].

(1) $\sum_{n\geq 1} (\theta-\lambda_n) < \infty$.

If $P < \sum_{n\geq 1} (\theta-\lambda_n)$, then there exists finite J such that the optimum no-feedback covariance is given by [4, Theorem 3],

$$R_X = \frac{1}{J}\left[\sum_{i=1}^{J} \lambda_i + P\right]\sum_{n=1}^{J} R_W^{\frac{1}{2}}e_n \otimes R_W^{\frac{1}{2}}e_n - R_W^{\frac{1}{2}}\left[\sum_{n=1}^{J} \lambda_n e_n \otimes e_n\right]R_W^{\frac{1}{2}}. \qquad (*)$$

As in the finite-dimensional channel, this shows that VR_XV^{*-1} will not be diagonal if the subspace spanned by the eigenvectors of $V^{-1}R_NV^{*-1}$ corresponding to the sequence of eigenvalues $(1 + \lambda_k, k \leq J)$ of $V^{-1}R_NV^{*-1}$ does not contain a basis consisting entirely of natural basis vectors which are eigenvectors of $V^{-1}R_NV^{*-1}$.

If $P = \sum_{n=1}^{\infty} (\theta-\lambda_n)$, then $J\lambda_J \leq P + \sum_{i=1}^{J} \lambda_i$ for every λ_J [4], and an optimum message covariance exists and is given by

$$R_X = \sum_{n\geq 1} (\theta-\lambda_n)R_W^{\frac{1}{2}}e_n \otimes R_W^{\frac{1}{2}}e_n.$$

This gives

$$V^{-1}R_XV^{*-1} = \sum_{n\geq 1} (\theta-\lambda_n)(L^*e_n)\otimes(L^*e_n),$$

which is clearly diagonal if and only if $\{L^*e_n, n \geq 1\}$ can be taken as natural basis vectors.

If (λ_n) is an infinite sequence and $P > \Sigma_n(\theta-\lambda_n)$, then capacity cannot be attained in the no-feedback case. However, the capacity is given by $\lim_{K\to\infty} I_W^K(P)$,

where $I^K_W(P)$ is the value of $I(X^K, X^K+N)$ when

$$R^K_X = \frac{1}{K}\left[\sum_{i=1}^{K} \lambda_i + P\right] \sum_{n=1}^{K} R_W^{\frac{1}{2}} e_n \otimes R_W^{\frac{1}{2}} e_n - R_W^{\frac{1}{2}}\left[\sum_{n=1}^{K} \lambda_n e_n \otimes e_n\right] R_W^{\frac{1}{2}}.$$

Let $\Delta_K(1)$ be the squared $H(\ell_2, \mu)$ norm of the projection of $V^{-1}X^K$ onto $H_-(X^K+N)$. If $\limsup_K \Delta_K(1) > 0$, then as before the capacity can be increased. That is, we choose K sufficiently large so that $I(\alpha X^K, \alpha X^K+N) > C_W(P)$, where $\alpha > 1$, $\alpha \geq \alpha_K(1)$, with $\alpha_K(1)$ satisfying $\alpha_K^2 = \dfrac{P - \mathrm{Tr}\, D_K R_N D_K}{P - \mathrm{Tr}\, D_K R_N D_K^* - \Delta_K(1)}$, with D_K the bounded SLT matrix that minimizes $E\|V^{-1}X^K - D_K(X^K+N)\|^2$. The problem is now reduced to showing that $\Delta_K(1) \to 0$ cannot hold if there exists some J such that $\{L^* e_n, n \leq J\}$ cannot be taken to be natural basis vectors. Suppose such J exists and take $K > J$.

Write $X^K = X^J_{OK} + X_{OK}$, where X^J_{OK} is the zero-mean Gaussian process with covariance matrix $R^{K,J}_X$,

$$R^{K,J}_X = \frac{1}{K} \sum_{n=1}^{J} \left[\sum_{i=1}^{K} \lambda_i + P - K\lambda_n\right] R_W^{\frac{1}{2}} e_n \otimes R_W^{\frac{1}{2}} e_n,$$

and X_{OK} is independent of X^J_{OK}. As $K \to \infty$, $R^{K,J}_X$ converges in the operator norm topology to $\sum_{n=1}^{J} (\theta-\lambda_n) R_W^{\frac{1}{2}} e_n \otimes R_W^{\frac{1}{2}} e_n$, using the fact that $\sum_{n \geq 1} (\theta-\lambda_n) < P$. Now suppose that $\Delta_K(1) \to 0$. This requires that $E\langle V^{-1}X^J_{OK} + V^{-1}X_{OK}, B(X^J_{OK}+X_{OK}+N)\rangle \to 0$ for every fixed bounded SLT matrix B. Since X^J_{OK} and X_{OK} are independent, this implies that $E\langle V^{-1}X^J_{OK}, BX^J_{OK}\rangle \to 0$ for every bounded SLT B, or $\{\mathrm{Trace}\, BR^{K,J}_X V^{*-1}\} \to 0$. Since $R^{K,J}_X \to \sum_{n=1}^{J} (\theta-\lambda_n) R_W^{\frac{1}{2}} e_n \otimes R_W^{\frac{1}{2}} e_n$ as $K \to \infty$, this implies (as in the proof of the Proposition) that $\sum_{n=1}^{J} (\theta-\lambda_n) L^* e_n \otimes L^* e_n$ is diagonal. This cannot be, since by assumption $\{L^* e_n, n \leq J\}$ cannot be taken as natural basis

vectors. This shows that optimum feedback will increase capacity when $P > \sum_{n\geq 1}(\theta-\lambda_n)$ and (λ_n) is an infinite sequence.

Finally, suppose that (λ_n) is a finite sequence, $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_K$, and $P > \sum_{n=1}^{K}(\theta-\lambda_n)$. Then there exists an infinite o.n. set $\{u_n, n\geq 1\}$ such that $\|(S-\theta I)u_n\| \to 0$ and $u_n \perp \text{span}\{e_1,\ldots,e_K\}$ for all $n \geq 1$ [6].

Fix $M < \infty$ and take $\epsilon > 0$ such that $\theta - \lambda_K > \epsilon$.

Let X_1^M be the zero-mean Gaussian process with covariance matrix R_1^M given by

$$R_1^M = \sum_{n=1}^{K}\left[\frac{\sum_{i=1}^{K}\lambda_i + P_1^M - K\lambda_n}{K(1+\lambda_n)}\right]R_N^{\frac{1}{2}}Ue_n \otimes R_N^{\frac{1}{2}}Ue_n$$

where $P_1^M < P$.

Choose o.n. vectors $u_1^{M,\epsilon},\ldots,u_2^{M,\epsilon}$ from the set $\{u_n, n \geq 1\}$ such that $|\langle(S-\theta I)u_i^{M,\epsilon}, u_i^{M,\epsilon}\rangle| \leq \epsilon$ for $i \leq M$. Let $X_2^{M,\epsilon}$ be the zero-mean Gaussian process with covariance matrix $R_2^{M,\epsilon}$ given by

$$R_2^{M,\epsilon} = \frac{P - P_1^M}{M(1+\theta+\epsilon)}\sum_{i=1}^{M} R_N^{\frac{1}{2}}Uu_i^{M,\epsilon}\otimes R_N^{\frac{1}{2}}Uu_i^{M,\epsilon}.$$

Now, let $X^{M,\epsilon}$ be the zero-mean Gaussian process with covariance matrix $R_X^{M,\epsilon}$, where $R_X^{M,\epsilon} = R_1^M + R_2^{M,\epsilon}$. Since $u_i^{M,\epsilon}$ is orthogonal to $\text{span}\{e_1,\ldots,e_K\}$ for $i \leq M$,

$$\begin{aligned} I(X^{M,\epsilon}, X^{M,\epsilon}+N) &= I(X_1^M, X_1^M+N) + I(X_2^{M,\epsilon}, X_2^{M,\epsilon}+N) \\ &= \tfrac{1}{2}\sum_{n=1}^{K}\log\left[1 + \frac{\sum_{i=1}^{K}\lambda_i + P_1^M - K\lambda_n}{K(1+\lambda_n)}\right] + \tfrac{1}{2}M\log\left[1 + \frac{P - P_1^M}{M(1+\theta+\epsilon)}\right]. \end{aligned}$$

$X^{M,\epsilon}$ satisfies the constraint for any $P_1^M < P$, since

$$E\|\underset{\sim}{X}^{M,\epsilon}\|_W^2 = \text{Trace } R_W^{-\frac{1}{2}} R_X^{M,\epsilon} R_W^{-\frac{1}{2}} = \text{Trace } R_W^{-\frac{1}{2}}(R_1^M + R_2^{M,\epsilon}) R_W^{-\frac{1}{2}}$$

$$\leq P_1^M + \frac{P - P_1^M}{M(1+\theta+\epsilon)} M(1 + \theta + \epsilon) = P.$$

Now, define P_1^M by

$$P_1^M = M^{-1}\left[KP - (M-K)\sum_{i=1}^{K}\lambda_i + (M-K)K\theta\right].$$

Then

$$I(X_1^M, X_1^M+N) = \frac{1}{2}\sum_{n=1}^{K}\log\left[\frac{MK + K\Sigma_{i=1}^K\lambda_i + KP + (M-K)K\theta}{MK(1+\lambda_n)}\right].$$

As $M \to \infty$,

$$I(X_1^M, X_1^M+N) \to \frac{1}{2}\sum_{n=1}^{K}\log\left[\frac{1+\theta}{1+\lambda_n}\right].$$

Similarly,

$$I(X_2^{M,\epsilon}, X_2^{M,\epsilon}+N) = \frac{1}{2} M \log\left[\frac{M^2(1+\theta+\epsilon) + (M-K)[P + \Sigma_{i=1}^K\lambda_i - K\theta]}{M^2(1+\theta+\epsilon)}\right].$$

As $M \to \infty$,

$$I(X_2^{M,\epsilon}, X_2^{M,\epsilon}+N) \longrightarrow \frac{P + \Sigma_{i=1}^K(\lambda_i - \theta)}{2(1+\theta+\epsilon)} .$$

Thus, $I(X^{M,\epsilon}, X^{M,\epsilon}+N)$ converges, as $M \to \infty$, to

$$\frac{1}{2}\sum_{n=1}^{K}\log\left[\frac{1+\theta}{1+\lambda_n}\right] + \frac{1}{2}\,\frac{P + \Sigma_{i=1}^K(\lambda_i - \theta)}{1 + \theta + \epsilon} .$$

From Theorem 3 of [4], the capacity $C_W(P)$ is equal to

$$\frac{1}{2}\sum_{n=1}^{K}\log\left[\frac{1+\theta}{1+\lambda_n}\right] + \frac{1}{2}\,\frac{P + \Sigma_{i=1}^K(\lambda_i - \theta)}{1 + \theta} .$$

so that by choosing ϵ sufficiently small and M sufficiently large, one can obtain $I(X^{M,\epsilon}, X^{M,\epsilon}+N)$ arbitrarily near $C_W(P)$.

As $\epsilon \to 0$ and $M \to \infty$, $V^{-1}R_2^{M,\epsilon}V^{*-1}$ converges to a diagonal matrix, since $|\langle(S-\theta I)u_i^{M,\epsilon}, u_i^{M,\epsilon}\rangle| \leq \epsilon$ for $i \leq M$. However,

$$V^{-1}R_1^M V^{*-1} = \sum_{n=1}^{K}\left[\frac{\Sigma_{i=1}^K \lambda_i + P_1^M - K\lambda_n}{K}\right]L^*e_n \otimes L^*e_n.$$

This matrix will not be diagonal if $\{L^*e_n, n \leq K\}$ cannot be taken to consist entirely of natural basis vectors. This is equivalent to not having K natural basis vectors as eigenvectors of $V^{-1}R_N V^{*-1}$ corresponding to the sequence of eigenvalues $(1 + \lambda_k, k \leq K)$. Inserting the above definition of P_1^M, $V^{-1}R_1^M V^{*-1} = \sum_{n=1}^{K}\left[(\theta-\lambda_n) + \left[P + \sum_{j=1}^{K}\lambda_j - K\theta\right]M^{-1}\right]L^*e_n \otimes L^*e_n$. This matrix is independent of ϵ; as $M \to \infty$, it converges in the operator norm topology to $V^{-1}R_1V^{*-1} = \sum_{n=1}^{K}(\theta-\lambda_n)L^*e_n \otimes L^*e_n$. Similar to the preceding part of the proof, we now consider $\Delta_{M,\epsilon}(1)$, the squared $H(\ell_2,\mu)$ norm of the projection of $V^{-1}X^{M,\epsilon}$ onto $H\_(X^{M,\epsilon}+N)$. $\Delta_{M,\epsilon}(1) \to 0$ as $M \to \infty$ implies, as in the preceding part of the proof, that $V^{-1}R_1V^{*-1}$ is diagonal. This is a contradiction. In fact, $\Delta_{M,\epsilon}(1)$ is bounded away from zero. Define $\alpha^2_{M,\epsilon} = \dfrac{P - \mathrm{Tr}\, D_{M,\epsilon}R_N D_{M,\epsilon}}{P - \mathrm{Tr}\, D_{M,\epsilon}R_N D_{M,\epsilon} - \Delta_{M,\epsilon}(1)}$ with $D_{M,\epsilon}(X^{M,\epsilon}+N)$ the projection of $V^{-1}X^{M,\epsilon}$ onto $H\_(X^{M,\epsilon}+N)$. Since $I(X^{M,\epsilon}, X^{M,\epsilon}+N) \to C_W(P)$ as $\epsilon \to 0$, $M \to \infty$, and $\Delta_{M,\epsilon}(1)$ is bounded away from zero, we obtain $I(\alpha_{M,\epsilon}X^{M,\epsilon}, \alpha_{M,\epsilon}X^{M,\epsilon}+N) > C_W(P)$. This completes the proof of sufficiency in part (2) of the Theorem when $\Sigma_n(\theta-\lambda_n) < \infty$.

(2) $\Sigma_n(\theta-\lambda_n) = \infty$.

In this case, $P < \Sigma_n(\theta-\lambda_n)$ for all $P > 0$, capacity is attained in the no-feedback case for every $P > 0$, and for each $P > 0$ there exists $J < \infty$ (the

value of J depending on P) such that the optimum message covariance matrix R_X is given as in (*). As in case (1), it is clear that feedback can increase capacity if the set of eigenvectors corresponding to the sequence of eigenvalues $(1 + \lambda_k,\ k \leq J)$ of $V^{-1}R_N V^{*-1}$ does not contain J natural basis vectors.

This completes the proof of (2) of the Theorem. The proof of (2) of the Corollary follows from (2) of the Theorem, in the same way that (1) of the Corollary was obtained. □

Verification of the Sufficient Conditions

Verification of the sufficient conditions given in the Theorem is equivalent to determining the value of $C_W(P)$, as can be seen from the expressions for $C_W(P)$ [4]. The difficulty of verifying the sufficient conditions of the Corollary is considerably less than for the Theorem. We now summarize how one can verify that $C_{WF}(P) > C_W(P)$ for all $P > 0$. This will be done by giving conditions that are equivalent to the conditions given in (1) and (2) of the Corollary for $C_{WF}(P) > C_W(P)$ for all $P > 0$.

Suppose that $V^{-1}R_N V^{*-1}$ is nondiagonal. Write

$$V^{-1}R_N V^{*-1} = A - D$$

where D is a diagonal matrix whose non-zero elements D(i,i) are the diagonal elements γ_{ii} of $V^{-1}R_N V^{*-1}$ such that $(V^{-1}R_N V^{*-1})(ij) = (V^{-1}R_N V^{*-1})(ji) = 0$ for all $j \neq i$. $C_{WF}(P) > C_W(P)$ for all $P > 0$ if the following conditions are satisfied.

1. Finite-Dimensional Channel.

$$\inf_{\|x\|=1} \langle Ax, x\rangle \leq \inf \{D(i,i) : D(i,i) > 0\};$$

2. Infinite-Dimensional Channel.

(a) $\inf_{\|x\|=1} \langle Ax,x\rangle \leq \inf \{D(i,i) : D(i,i) > 0\}$;

(b) $\inf_{\|x\|=1} \langle Ax,x\rangle$ is an eigenvalue of $V^{-1}R_N V^{*-1}$ of finite multiplicity;

and

(c) if H_0 is the subspace spanned by the eigenvectors of $V^{-1}R_N V^{*-1}$ corresponding to the eigenvalue $\inf_{\|x\|=1} \langle Ax,x\rangle$, then

$$\inf_{\substack{\|x\|=1 \\ x \text{ in } H_0^{\perp}}} \langle Ax,x\rangle > \inf_{\|x\|=1} \langle Ax,x\rangle .$$

To see that (2) implies the corresponding sufficient condition in (2) of the Corollary, one can verify that (2a) and (2b) imply that the smallest eigenvalue of $V^{-1}R_N V^{*-1}$ exists and does not have eigenspace containing a set of natural basis vectors complete for the subspace; (2b) shows that the multiplicity of this subspace is finite; and (2b) plus (2c) show that this eigenvalue is not the limit of a sequence of distinct eigenvalues.

These conditions are not complex. Consider the finite-dimensional channel. First, one inspects the matrix $V^{-1}R_N V^{*-1}$ and locates the diagonal elements γ_{ii} such that the i<u>th</u> row and i<u>th</u> column are all zero except for the ii element. Denote these elements as γ_i. This is the set of eigenvalues of $V^{-1}R_N V^{*-1}$ corresponding to natural basis vectors as eigenvectors. If the smallest such γ_i is strictly greater than $\inf_{\|x\|=1} \langle V^{-1}R_N V^{*-1}x,x\rangle$, then the smallest eigenvalue of $V^{-1}R_N V^{*-1}$ has no eigenvectors that are natural basis vectors, and so $C_{WF}(P) > C_W(P)$ for all $P > 0$. If the smallest γ_i is equal to $\inf_{\|x\|=1} \langle V^{-1}R_N V^{*-1}x,x\rangle$, then one must determine the multiplicity of $\inf_i \gamma_i = \gamma_0$

as an eigenvalue of $V^{-1}R_N V^{*-1}$. If this multiplicity is strictly greater than the number of times γ_0 appears among the $\{\gamma_i, i \geq 1\}$, then again $C_{WF}(P) > C_W(P)$ for all $P > 0$.

Necessary Conditions

The Corollary shows that the sufficient condition for feedback to increase capacity for all sufficiently large P is also necessary, in the case of the finite-dimensional channel. Although the emphasis here has been on sufficient conditions, it is our conjecture that each of the four sufficient conditions given in the Corollary is also a necessary condition for the same result.

Concluding Remarks

It can be seen that the capacity problem with feedback for small P reduces to consideration of the eigenmanifold for the smallest eigenvalue of $V^{-1}R_N V^{*-1}$, for the finite-dimensional channel. If this eigenvalue has multiplicity one, then feedback can increase capacity for every value of P if the corresponding eigenvector is not a natural basis vector.

In the case of the infinite-dimensional channel, the same situation holds, except that the additional requirement is imposed of having the smallest eigenvalue be strictly less than the smallest limit point of $V^{-1}R_N V^{*-1}$.

For the case of sufficiently large P, the problem can be couched in terms of the reproducing kernel Hilbert space of R_W, say H_W. If the Gaussian cylinder set measure μ on H_W defined by $\mu_N = \mu \circ j^{-1}$, j the natural injection of H_W into ℓ_2 (i.e., jx is just x viewed as an element of ℓ_2 rather than as an element of H_W), has diagonal covariance operator, then $C_{WF}(P) = C(P)$; otherwise,

$C_{WF}(P) > C_W(P)$ for all sufficiently large P. In essence, this states that capacity can be increased by feedback for all sufficiently large P if the noise is correlated when it is viewed as belonging to H_W, rather than to ℓ_2.

The setup given here is rather general, and an obvious extension is to apply the same approach to the time-continuous channel. However, the structure of (Hilbert-Schmidt) Volterra operators is more complicated in $L_2[0,T]$ than in ℓ_2, and an arbitrary covariance operator in $L_2[0,T]$ may not have a causal decomposition of the form $R_W = VV^*$, V Volterra. Thus, a complete extension of these results in the form stated here does not seem possible.

References

1. C.R. Baker, Information and coding capacities of mismatched Gaussian channels, to appear, *Proceedings of International Conference on Advances in Communication and Control Systems*, Washington, D.C., June 18-20, 1987.

2. S. Ihara, Capacity of Discrete-Time Gaussian Channel with and without Feedback, preprint, 1987.

3. T.M. Cover and S. Pombra, Gaussian Feedback Capacity, preprint, 1987.

4. C.R. Baker, Capacity of the mismatched Gaussian channel, *IEEE Trans. on Inform. Theory*, 33, No. 6, 1987.

5. R. Schatten, Norm Ideals of Completely Continuous Operators, Springer-Verlag, New York, 1960.

6. F. Riesz and B. Sz.-Nagy, Functional Analysis, Ungar, New York, 1955.

7. D.K. Fadeev and V.N. Fadeeva, Computational Methods of Linear Algebra, W.H. Freeman and Co., London, 1963.

8. R.G. Douglas, On majorization, factorization, and range inclusion of operators on Hilbert space, *Proc. Amer. Math. Soc.*, 17, 413-415 (1966).

END
DATE
FILMED
4-88
DTIC